

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Gotschlich, Emil C.
- (ii) TITLE OF INVENTION: GLYCOSYLTRANSFERASES FOR BIOSYNTHESIS OF OLIGOSACCHARIDES, AND GENES ENCODING THEM
- (iii) NUMBER OF SEQUENCES: 12
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: Klauber & Jackson
 - (B) STREET: 411 Hackensack Avenue
 - (C) CITY: Hackensack
 - (D) STATE: New Jersey
 - (E) COUNTRY: USA
 - (F) ZIP: 07601
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE:
 - (C) CLASSIFICATION:
- (vi) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 08/312,387
 - (B) FILING DATE: September 26, 1994
 - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Jackson Esq., David A.
 - (B) REGISTRATION NUMBER: 26,742
 - (C) REFERENCE/DOCKET NUMBER: 600-1-095B
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 201 487-5800
 - (B) TELEFAX: 201 343-1684
 - (C) TELEX: 133521

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5859 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Neisseria gonorrhoeae

(B) STRAIN: F62

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 1..381
(C) GENE: glys (glycyl tRNA syntetase beta chain)

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 445..1491
(C) GENE: lgtA

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 2342..3262
(C) GENE: lgtC

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 3322..4335
(C) GENE: lgtD

(ix) FEATURE:

(A) NAME/KEY: CDS
(B) LOCATION: 4354..5196
(C) GENE: lgtE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

| | |
|--|-----|
| CTG CAG GCC GTC GCC GTA TTC AAA CAA CTG CCC GAA GCC GCC GCG CTC | 48 |
| Leu Gln Ala Val Ala Val Phe Lys Gln Leu Pro Glu Ala Ala Ala Leu | |
| 1 5 10 15 | |
| GCC GCC GCC AAC AAA CGC GTG CAA AAC CTG CTG AAA AAA GCC GAT GCC | 96 |
| Ala Ala Ala Asn Lys Arg Val Gln Asn Leu Leu Lys Lys Ala Asp Ala | |
| 20 25 30 | |
| GCG TTG GGC GAA GTC AAT GAA AGC CTG CTG CAA CAG GAC GAA GAA AAA | 144 |
| Ala Leu Gly Glu Val Asn Glu Ser Leu Leu Gln Gln Asp Glu Glu Lys | |
| 35 40 45 | |
| GCC CTG TAC GCT GCC GCG CAA GGT TTG CAG CCG AAA ATT GCC GCC GCC | 192 |
| Ala Leu Tyr Ala Ala Ala Gln Gly Leu Gln Pro Lys Ile Ala Ala Ala | |
| 50 55 60 | |
| GTC GCC GAA GGC AAT TTC CGA ACC GCC TTG TCC GAA CTG GCT TCC GTC | 240 |
| Val Ala Glu Gly Asn Phe Arg Thr Ala Leu Ser Glu Leu Ala Ser Val | |
| 65 70 75 80 | |
| AAG CCG CAG GTT GAT GCC TTC TTC GAC GGC GTG ATG GTG ATG GCG GAA | 288 |
| Lys Pro Gln Val Asp Ala Phe Phe Asp Gly Val Met Val Met Ala Glu | |
| 85 90 95 | |
| GAT GCC GCC GTA AAA CAA AAC CGC CTG AAC CTG CTG AAC CGC TTG GCA | 336 |
| Asp Ala Ala Val Lys Gln Asn Arg Leu Asn Leu Leu Asn Arg Leu Ala | |
| 100 105 110 | |
| GAG CAG ATG AAC GCG GTG GCC GAC ATC GCG CTT TTG GGC GAG TAACCGTTGT | 388 |
| Glu Gln Met Asn Ala Val Ala Ile Ala Leu Leu Gly Glu | |
| 115 120 125 | |

| | |
|---|------|
| ACAGTCCAAA TGCCGTCTGA AGCCTTCAGG CGGCATCAAA TTATCGGGAG AGTAAA | 444 |
| TTG CAG CCT TTA GTC AGC GTA TTG ATT TGC GCC TAC AAC GTA GAA AAA Met Gln Pro Leu Val Ser Val Leu Ile Cys Ala Tyr Asn Val Glu Lys 1 5 10 15 | 492 |
| TAT TTT GCC CAA TCA TTA GCC GCC GTC GTG AAT CAG ACT TGG CGC AAC Tyr Phe Ala Gln Ser Leu Ala Ala Val Val Asn Gln Thr Trp Arg Asn 20 25 30 | 540 |
| TTG GAT ATT TTG ATT GTC GAT GAC GGC TCG ACA GAC GGC ACA CTT GCC Leu Asp Ile Leu Ile Val Asp Asp Gly Ser Thr Asp Gly Thr Leu Ala 35 40 45 | 588 |
| ATT GCC AAG GAT TTT CAA AAG CGG GAC AGC CGT ATC AAA ATC CTT GCA Ile Ala Lys Asp Phe Gln Lys Arg Asp Ser Arg Ile Lys Ile Leu Ala 50 55 60 | 636 |
| CAA GCT CAA AAT TCC GGC CTG ATT CCC TCT TTA AAC ATC GGG CTG GAC Gln Ala Gln Asn Ser Gly Leu Ile Pro Ser Leu Asn Ile Gly Leu Asp 65 70 75 80 | 684 |
| GAA TTG GCA AAG TCG GGG GGG GGG GGG GAA TAT ATT GCG CGC ACC Glu Leu Ala Lys Ser Gly Gly Gly Gly Glu Tyr Ile Ala Arg Thr 85 90 95 | 732 |
| GAT GCC GAC GAT ATT GCC TCC CCC GGC TGG ATT GAG AAA ATC GTG GGC Asp Ala Asp Ile Ala Ser Pro Gly Trp Ile Glu Lys Ile Val Gly 100 105 110 | 780 |
| GAG ATG GAA AAA GAC CGC AGC ATC ATT GCG ATG GGC GCG TGG CTG GAA Glu Met Glu Lys Asp Arg Ser Ile Ile Ala Met Gly Ala Trp Leu Glu 115 120 125 | 828 |
| GTT TTG TCG GAA GAA AAG GAC GGC AAC CGG CTG GCG CGG CAC CAC AAA Val Leu Ser Glu Glu Lys Asp Gly Asn Arg Leu Ala Arg His His Lys 130 135 140 | 876 |
| CAC GGC AAA ATT TGG AAA AAG CCG ACC CGG CAC GAA GAC ATC GCC GCC His Gly Lys Ile Trp Lys Lys Pro Thr Arg His Glu Asp Ile Ala Ala 145 150 155 160 | 924 |
| TTT TTC CCT TTC GGC AAC CCC ATA CAC AAC AAC ACG ATG ATT ATG CGG Phe Phe Pro Phe Gly Asn Pro Ile His Asn Asn Thr Met Ile Met Arg 165 170 175 | 972 |
| CGC AGC GTC ATT GAC GGC GGT TTG CGT TAC GAC ACC GAG CGG GAT TGG Arg Ser Val Ile Asp Gly Gly Leu Arg Tyr Asp Thr Glu Arg Asp Trp 180 185 190 | 1020 |
| GCG GAA GAT TAC CAA TTT TGG TAC GAT GTC AGC AAA TTG GGC AGG CTG Ala Glu Asp Tyr Gln Phe Trp Tyr Asp Val Ser Lys Leu Gly Arg Leu 195 200 205 | 1068 |
| GCT TAT TAT CCC GAA GCC TTG GTC AAA TAC CGC CTT CAC GCC AAT CAG Ala Tyr Tyr Pro Glu Ala Leu Val Lys Tyr Arg Leu His Ala Asn Gln 210 215 220 | 1116 |
| GTT TCA TCC AAA CAC AGC GTC CGC CAA CAC GAA ATC GCG CAA GGC ATC Val Ser Ser Lys His Ser Val Arg Gln His Glu Ile Ala Gln Gly Ile 225 230 235 240 | 1164 |

| | |
|---|------|
| CAA AAA ACC GCC AGA AAC GAT TTT TTG CAG TCT ATG GGT TTT AAA ACC Gln Lys Thr Ala Arg Asn Asp Phe Leu Gln Ser Met Gly Phe Lys Thr 245 250 255 | 1212 |
| CGG TTC GAC AGC CTA GAA TAC CGC CAA ACA AAA GCA GCG GCG TAT GAA Arg Phe Asp Ser Leu Glu Tyr Arg Gln Thr Lys Ala Ala Ala Tyr Glu 260 265 270 | 1260 |
| CTG CCG GAG AAG GAT TTG CCG GAA GAA GAT TTT GAA CGC GCC CGC CGG Leu Pro Glu Lys Asp Leu Pro Glu Glu Asp Phe Glu Arg Ala Arg Arg 275 280 285 | 1308 |
| TTT TTG TAC CAA TGC TTC AAA CGG ACG GAC ACG CCG CCC TCC GGC GCG Phe Leu Tyr Gln Cys Phe Lys Arg Thr Asp Thr Pro Pro Ser Gly Ala 290 295 300 | 1356 |
| TGG CTG GAT TTC GCG GCA GAC GGC AGG ATG AGG CGG CTG TTT ACC TTG Trp Leu Asp Phe Ala Ala Asp Gly Arg Met Arg Arg Leu Phe Thr Leu 305 310 315 320 | 1404 |
| AGG CAA TAC TTC GGC ATT TTG TAC CGG CTG ATT AAA AAC CGC CGG CAG Arg Gln Tyr Phe Gly Ile Leu Tyr Arg Leu Ile Lys Asn Arg Arg Gln 325 330 335 | 1452 |
| GCG CGG TCG GAT TCG GCA GGG AAA GAA CAG GAG ATT TAATGCAAAA Ala Arg Ser Asp Ser Ala Gly Lys Glu Gln Glu Ile 340 345 | 1498 |
| CCACGTTATC AGCTTGGCTT CCGCCGCAGA ACGCAGGGCG CACATTGCCG CAACCTTCGG | 1558 |
| CAGTCGCGGC ATCCCGTTCC AGTTTTTCGA CGCACTGATG CCGTCTGAAA GGCTGGAACG | 1618 |
| GGCAATGGCG GAACTCGTCC CCGGCTTGTC GGCGCACCCC TATTTGAGCG GAGTGGAAAA | 1678 |
| AGCCTGCTTT ATGAGCCACG CCGTATTGTG GGAACAGGCA TTGGACGAAG GCGTACCGTA | 1738 |
| TATCGCCGTA TTTGAAGATG ATGTCTTACT CGGCGAAGGC GCGGAGCAGT TCCTTGCCGA | 1798 |
| AGATACTTGG CTGCAAGAAC GCTTTGACCC CGATTCCGCC TTTGTCGTCC GCTTGGAAC | 1858 |
| GATGTTTATG CACGTCCTGA CCTCGCCCTC CGGCGTGGCG GACTACGGCG GGCGCGCCTT | 1918 |
| TCCGCTTTTG GAAAGCGAAC ACTGCGGGAC GGCGGGCTAT ATTATTTCCC GAAAGGCGAT | 1978 |
| GCGTTTTTTC TTGGACAGGT TTGCCGTTT GCCGCCCGAA CGCCTGCACC CTGTGATTT | 2038 |
| GATGATGTTT GGCAACCTTG ACGACAGGGA AGGAATGCCG GTTTGCCAGC TCAATCCCGC | 2098 |
| CTTGTGCGCC CAAGAGCTGC ATTATGCCAA GTTTCACGAC CAAAACAGCG CATTGGGCAG | 2158 |
| CCTGATCGAA CATGACCGCC GCCTGAACCG CAAACAGCAA TGGCGCGATT CCCCCGCCAA | 2218 |
| CACATTCAAA CACCGCCTGA TCCGCGCCTT GACCAAAATC GGCAGGGAAA GGGAAAAACG | 2278 |
| CCGGCAAAGG CGCGAACAGT TAATCGGCAA GATTATTGTG CCTTTCCAAT AAAAGGAGAA | 2338 |
| AAG ATG GAC ATC GTA TTT GCG GCA GAC GAC AAC TAT GCC GCC TAC CTT Met Asp Ile Val Phe Ala Ala Asp Asp Asn Tyr Ala Ala Tyr Leu 1 5 10 15 | 2386 |
| TGC GTT GCG GCA AAA AGC GTG GAA GCG GCC CAT CCC GAT ACG GAA ATC Cys Val Ala Ala Lys Ser Val Glu Ala Ala His Pro Asp Thr Glu Ile | 2434 |

| 20 | | | | | | | | | | 25 | | | | 30 | | | | |
|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|------|--|--|
| AGG Arg | TTC Phe | CAC His | GTC Val 35 | CTC Leu | GAT Asp | GCC Ala | GGC Gly | ATC Ile 40 | AGT Ser | GAG Glu | GAA Glu | AAC Asn | CGG Arg 45 | GCG Ala | GCG Ala | 2482 | | |
| GTT Val | GCC Ala 50 | GCC Ala | AAT Asn | TTG Leu | CGG Arg | GGG Gly 55 | GGG Gly | GGT Gly | AAT Asn | ATC Ile | CGC Arg | TTT Phe 60 | ATA Ile | GAC Asp | GTA Val | 2530 | | |
| AAC Asn 65 | CCC Pro | GAA Glu | GAT Asp | TTC Phe | GCC Ala | GGC Gly 70 | TTC Phe | CCC Pro | TTA Leu | AAC Asn | ATC Ile 75 | AGG Arg | CAC His | ATT Ile | TCC Ser | 2578 | | |
| ATT Ile 80 | ACG Thr | ACT Thr | TAT Tyr | GCC Ala | CGC Arg 85 | CTG Leu | AAA Lys | TTG Leu | GGC Gly | GAA Glu 90 | TAC Tyr | ATT Ile | GCC Ala | GAT Asp | TGC Cys 95 | 2626 | | |
| GAC Asp | AAA Lys | GTC Val | CTG Leu 100 | TAT Tyr | CTG Leu | GAT Asp | ACG Thr | GAC Asp 105 | GTA Val | TTG Leu | GTC Val | AGG Arg | GAC Asp | GGC Gly 110 | CTG Leu | 2674 | | |
| AAG Lys | CCC Pro | TTA Leu | TGG Trp 115 | GAT Asp | ACC Thr | GAT Asp | TTG Leu | GGC Gly 120 | GGT Gly | AAC Asn | TGG Trp | GTC Val | GGC Gly 125 | GCG Ala | TGC Cys | 2722 | | |
| ATC Ile | GAT Asp 130 | TTG Leu | TTT Phe | GTC Val | GAA Glu | AGG Arg 135 | CAG Gln | GAA Glu | GGA Gly | TAC Tyr | AAA Lys | CAA Gln 140 | AAA Lys | ATC Ile | GGT Gly | 2770 | | |
| ATG Met | GCG Ala 145 | GAC Asp | GGA Gly | GAA Glu | TAT Tyr 150 | TAT Tyr | TTC Phe | AAT Asn | GCC Ala | GGC Gly | GTA Val 155 | TTG Leu | CTG Leu | ATC Ile | AAC Asn | 2818 | | |
| CTG Leu 160 | AAA Lys | AAG Lys | TGG Trp | CGG Arg 165 | CGG Arg | CAC His | GAT Asp | ATT Ile | TTC Phe | AAA Lys 170 | ATG Met | TCC Ser | TGC Cys | GAA Glu | TGG Trp 175 | 2866 | | |
| GTG Val | GAA Glu | CAA Gln | TAC Tyr 180 | AAG Lys | GAC Asp | GTG Val | ATG Met | CAA Gln 185 | TAT Tyr | CAG Gln | GAT Asp | CAG Gln | GAC Asp | ATT Ile 190 | TTG Leu | 2914 | | |
| AAC Asn | GGG Gly | CTG Leu | TTT Phe 195 | AAA Lys | GGC Gly | GGG Gly | GTG Val | TGT Cys 200 | TAT Tyr | GCG Ala | AAC Asn | AGC Ser | CGT Arg 205 | TTC Phe | AAC Asn | 2962 | | |
| TTT Phe | ATG Met 210 | CCG Pro | ACC Thr | AAT Asn | TAT Tyr | GCC Ala | TTT Phe 215 | ATG Met | GCG Ala | AAC Asn | GGG Gly | TTT Phe 220 | GCG Ala | TCC Ser | CGC Arg | 3010 | | |
| CAT His | ACC Thr 225 | GAC Asp | CCG Pro | CTT Leu | TAC Tyr | CTC Leu 230 | GAC Asp | CGT Arg | ACC Thr | AAT Asn | ACG Thr 235 | GCG Ala | ATG Met | CCC Pro | GTC Val | 3058 | | |
| GCC Ala 240 | GTC Val | AGC Ser | CAT His | TAT Tyr | TGC Cys 245 | GGC Gly | TCG Ser | GCA Ala | AAG Lys | CCG Pro 250 | TGG Trp | CAC His | AGG Arg | GAC Asp | TGC Cys 255 | 3106 | | |
| ACC Thr | GTT Val | TGG Trp | GGT Gly 260 | GCG Ala | GAA Glu | CGT Arg | TTC Phe | ACA Thr 265 | GAG Glu | TTG Leu | GCC Ala | GGC Gly | AGC Ser | CTG Leu 270 | ACG Thr | 3154 | | |

| | |
|---|------|
| ACC GTT CCC GAA GAA TGG CGC GGC AAA CTT GCC GTC CCG CCG ACA AAG Thr Val Pro Glu Glu Trp Arg Gly Lys Leu Ala Val Pro Pro Thr Lys 275 280 285 | 3202 |
| TGT ATG CTT CAA AGA TGG CGC AAA AAG CTG TCT GCC AGA TTC TTA CGC Cys Met Leu Gln Arg Trp Arg Lys Lys Leu Ser Ala Arg Phe Leu Arg 290 295 300 | 3250 |
| AAG ATT TAT TGACGGGGCA GGCCGTCTGA AGCCTTCAGA CGGCATCGGA Lys Ile Tyr 305 | 3299 |
| CGTATCGGAA AGGAGAAACG GA TTG CAG CCT TTA GTC AGC GTA TTG ATT TGC Met Gln Pro Leu Val Ser Val Leu Ile Cys 1 5 10 | 3351 |
| GCC TAC AAC GCA GAA AAA TAT TTT GCC CAA TCA TTG GCC GCC GTA GTG Ala Tyr Asn Ala Glu Lys Tyr Phe Ala Gln Ser Leu Ala Ala Val Val 15 20 25 | 3399 |
| GGG CAG ACT TGG CGC AAC TTG GAT ATT TTG ATT GTC GAT GAC GGC TCG Gly Gln Thr Trp Arg Asn Leu Asp Ile Leu Ile Val Asp Asp Gly Ser 30 35 40 | 3447 |
| ACG GAC GGC ACG CCC GCC ATT GCC CGG CAT TTC CAA GAA CAG GAC GGC Thr Asp Gly Thr Pro Ala Ile Ala Arg His Phe Gln Glu Gln Asp Gly 45 50 55 | 3495 |
| AGG ATC AGG ATA ATT TCC AAT CCC CGC AAT TTG GGC TTT ATC GCC TCT Arg Ile Arg Ile Ile Ser Asn Pro Arg Asn Leu Gly Phe Ile Ala Ser 60 65 70 | 3543 |
| TTA AAC ATC GGG CTG GAC GAA TTG GCA AAG TCG GGG GGG GGG GAA TAT Leu Asn Ile Gly Leu Asp Glu Leu Ala Lys Ser Gly Gly Gly Glu Tyr 75 80 85 90 | 3591 |
| ATT GCG CGC ACC GAT GCC GAC GAT ATT GCC TCC CCC GGC TGG ATT GAG Ile Ala Arg Thr Asp Ala Asp Asp Ile Ala Ser Pro Gly Trp Ile Glu 95 100 105 | 3639 |
| AAA ATC GTG GGC GAG ATG GAA AAA GAC CGC AGC ATC ATT GCG ATG GGC Lys Ile Val Gly Glu Met Glu Lys Asp Arg Ser Ile Ile Ala Met Gly 110 115 120 | 3687 |
| GCG TGG TTG GAA GTT TTG TCG GAA GAA AAC AAT AAA AGC GTG CTT GCC Ala Trp Leu Glu Val Leu Ser Glu Glu Asn Asn Lys Ser Val Leu Ala 125 130 135 | 3735 |
| GCC ATT GCC CGA AAC GGC GCA ATT TGG GAC AAA CCG ACC CGG CAT GAA Ala Ile Ala Arg Asn Gly Ala Ile Trp Asp Lys Pro Thr Arg His Glu 140 145 150 | 3783 |
| GAC ATT GTC GCC GTT TTC CCT TTC GGC AAC CCC ATA CAC AAC AAC ACG Asp Ile Val Ala Val Phe Pro Phe Gly Asn Pro Ile His Asn Asn Thr 155 160 165 170 | 3831 |
| ATG ATT ATG AGG CGC AGC GTC ATT GAC GGC GGT TTG CGG TTC GAT CCA Met Ile Met Arg Arg Ser Val Ile Asp Gly Leu Arg Phe Asp Pro 175 180 185 | 3879 |
| GCC TAT ATC CAC GCC GAA GAC TAT AAG TTT TGG TAC GAA GCC GGC AAA Ala Tyr Ile His Ala Glu Asp Tyr Lys Phe Trp Tyr Glu Ala Gly Lys | 3927 |

| 190 | 195 | 200 | |
|---|-----|-----|------|
| CTG GGC AGG CTG GCT TAT TAT CCC GAA GCC TTG GTC AAA TAC CGC TTC Leu Gly Arg Leu Ala Tyr Tyr Pro Glu Ala Leu Val Lys Tyr Arg Phe 205 210 215 | | | 3975 |
| CAT CAA GAC CAG ACT TCT TCC AAA TAC AAC CTG CAA CAG CGC AGG ACG His Gln Asp Gln Thr Ser Ser Lys Tyr Asn Leu Gln Gln Arg Arg Thr 220 225 230 | | | 4023 |
| GCG TGG AAA ATC AAA GAA GAA ATC AGG GCG GGG TAT TGG AAG GCG GCA Ala Trp Lys Ile Lys Glu Glu Ile Arg Ala Gly Tyr Trp Lys Ala Ala 235 240 245 250 | | | 4071 |
| GGC ATA GCC GTC GGG GCG GAC TGC CTG AAT TAC GGG CTT TTG AAA TCA Gly Ile Ala Val Gly Ala Asp Cys Leu Asn Tyr Gly Leu Leu Lys Ser 255 260 265 | | | 4119 |
| ACG GCA TAT GCG TTG TAC GAA AAA GCC TTG TCC GGA CAG GAT ATC GGA Thr Ala Tyr Ala Leu Tyr Glu Lys Ala Leu Ser Gly Gln Asp Ile Gly 270 275 280 | | | 4167 |
| TGC CTC CGC CTG TTC CTG TAC GAA TAT TTC TTG TCG TTG GAA AAG TAT Cys Leu Arg Leu Phe Leu Tyr Glu Tyr Phe Leu Ser Leu Glu Lys Tyr 285 290 295 | | | 4215 |
| TCT TTG ACC GAT TTG CTG GAT TTC TTG ACA GAC CGC GTG ATG AGG AAG Ser Leu Thr Asp Leu Leu Asp Phe Leu Thr Asp Arg Val Met Arg Lys 300 305 310 | | | 4263 |
| CTG TTT GCC GCA CCG CAA TAT AGG AAA ATC CTG AAA AAA ATG TTA CGC Leu Phe Ala Ala Pro Gln Tyr Arg Lys Ile Leu Lys Lys Met Leu Arg 315 320 325 330 | | | 4311 |
| CCT TGG AAA TAC CGC AGC TAT TGAAACCGAA CAGGATAAAT C ATG CAA AAC Pro Trp Lys Tyr Arg Ser Tyr Met Gln Asn 335 1 | | | 4362 |
| CAC GTT ATC AGC TTG GCT TCC GCC GCA GAG CGC AGG GCG CAC ATT GCC His Val Ile Ser Leu Ala Ser Ala Ala Glu Arg Arg Ala His Ile Ala 5 10 15 | | | 4410 |
| GAT ACC TTC GGC AGT CGC GGC ATC CCG TTC CAG TTT TTC GAC GCA CTG Asp Thr Phe Gly Ser Arg Gly Ile Pro Phe Gln Phe Phe Asp Ala Leu 20 25 30 35 | | | 4458 |
| ATG CCG TCT GAA AGG CTG GAA CAG GCG ATG GCG GAA CTC GTC CCC GGC Met Pro Ser Glu Arg Leu Glu Gln Ala Met Ala Glu Leu Val Pro Gly 40 45 50 | | | 4506 |
| TTG TCG GCG CAC CCC TAT TTG AGC GGA GTG GAA AAA GCC TGC TTT ATG Leu Ser Ala His Pro Tyr Leu Ser Gly Val Glu Lys Ala Cys Phe Met 55 60 65 | | | 4554 |
| AGC CAC GCC GTA TTG TGG GAA CAG GCG TTG GAT GAA GGT CTG CCG TAT Ser His Ala Val Leu Trp Glu Gln Ala Leu Asp Glu Gly Leu Pro Tyr 70 75 80 | | | 4602 |
| ATC GCC GTA TTT GAG GAC GAC GTT TTA CTC GGC GAA GGC GCG GAG CAG Ile Ala Val Phe Glu Asp Asp Val Leu Leu Gly Glu Gly Ala Glu Gln 85 90 95 | | | 4650 |

| | |
|---|------|
| TTC CTT GCC GAA GAT ACT TGG TTG GAA GAG CGT TTT GAC AAG GAT TCC Phe Leu Ala Glu Asp Thr Trp Leu Glu Glu Arg Phe Asp Lys Asp Ser 100 105 110 115 | 4698 |
| GCC TTT ATC GTC CGT TTG GAA ACG ATG TTT GCG AAA GTT ATT GTC AGA Ala Phe Ile Val Arg Leu Glu Thr Met Phe Ala Lys Val Ile Val Arg 120 125 130 | 4746 |
| CCG GAT AAA GTC CTG AAT TAT GAA AAC CGG TCA TTT CCT TTG CTG GAG Pro Asp Lys Val Leu Asn Tyr Glu Asn Arg Ser Phe Pro Leu Leu Glu 135 140 145 | 4794 |
| AGC GAA CAT TGT GGG ACG GCT GGC TAT ATC ATT TCG CGT GAG GCG ATG Ser Glu His Cys Gly Thr Ala Gly Tyr Ile Ile Ser Arg Glu Ala Met 150 155 160 | 4842 |
| CGG TTT TTC TTG GAC AGG TTT GCC GTT TTG CCG CCA GAG CGG ATT AAA Arg Phe Phe Leu Asp Arg Phe Ala Val Leu Pro Pro Glu Arg Ile Lys 165 170 175 | 4890 |
| GCG GTA GAT TTG ATG ATG TTT ACT TAT TTC TTT GAT AAG GAG GGG ATG Ala Val Asp Leu Met Met Phe Thr Tyr Phe Phe Asp Lys Glu Gly Met 180 185 190 195 | 4938 |
| CCT GTT TAT CAG GTT AGT CCC GCC TTA TGT ACC CAA GAA TTG CAT TAT Pro Val Tyr Gln Val Ser Pro Ala Leu Cys Thr Gln Glu Leu His Tyr 200 205 210 | 4986 |
| GCC AAG TTT CTC AGT CAA AAC AGT ATG TTG GGT AGC GAT TTG GAA AAA Ala Lys Phe Leu Ser Gln Asn Ser Met Leu Gly Ser Asp Leu Glu Lys 215 220 225 | 5034 |
| GAT AGG GAA CAA GGA AGA AGA CAC CGC CGT TCG TTG AAG GTG ATG TTT Asp Arg Glu Gln Gly Arg Arg His Arg Arg Ser Leu Lys Val Met Phe 230 235 240 | 5082 |
| GAC TTG AAG CGT GCT TTG GGT AAA TTC GGT AGG GAA AAG AAG AAA AGA Asp Leu Lys Arg Ala Leu Gly Lys Phe Gly Arg Glu Lys Lys Lys Arg 245 250 255 | 5130 |
| ATG GAG CGT CAA AGG CAG GCG GAG CTT GAG AAA GTT TAC GGC AGG CGG Met Glu Arg Gln Arg Gln Ala Glu Leu Glu Lys Val Tyr Gly Arg Arg 260 265 270 275 | 5178 |
| GTC ATA TTG TTC AAA TAGTTTGTGT AAAATATAGG GGATTAAAAT CAGAAATGGA Val Ile Leu Phe Lys 280 | 5233 |
| CACACTGTCA TTCCCGCGCA GCGGGAATC TAGGTCTTTA AACTTCGGTT TTTTCCGATA | 5293 |
| AATTCCTTGCC GCATTAAAAT TCCAGATTCC CGCTTTCGCG GGGATGACGG CGGGGGGATT | 5353 |
| GTGCTTTTTT CGGATAAAAT CCCGTGTTTT TTCATCTGCT AGGTAAAATC GCCCAAAGC | 5413 |
| GTCTGCATCG CGGCGATGGC GCGGAGTGGG GCGGTTTCTG TCGTAAAAT CCGTTTTCCG | 5473 |
| AGTGTAACCG CCTGAAAGCC GGCTTCAAAT GCCTGTTGTT CTTCTGTTC TGTCCAGCCG | 5533 |
| CCTTCGGGCC CGACCATAAA GACGATTGCG CCGGACGGGT GCGGATGTC GCCGAGTTTG | 5593 |
| CAGGCGCGGT TGATGCTCAT AATCAGCTTG GTGTTTTAG ACGGCATTTT GTCGAGTGCT | 5653 |

| | |
|--|------|
| TCACGGTAGC CGATGATGGG CAGTACGGGG GGAACGGTGT TCCTGCCGCT TTGTTGCGAC | 5713 |
| GCGGAGATGA CGATTTTCCTG CCAGCGTGCG AGGCGTTTGG CGGCGCGTTC TCCGTCGAGG | 5773 |
| CGGACGATGC AGCGTTCGCT GATGACGGGC TGTATGGCGG TTACGCCGAG TTCGACGCTT | 5833 |
| TTTTGCAGGG TGAAATCCAT GCGATC | 5859 |

(2) INFORMATION FOR SEQ ID NO:2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 126 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|
| Leu | Gln | Ala | Val | Ala | Val | Phe | Lys | Gln | Leu | Pro | Glu | Ala | Ala | Ala | Leu | 1 | 5 | 10 | 15 |
| Ala | Ala | Ala | Asn | Lys | Arg | Val | Gln | Asn | Leu | Leu | Lys | Lys | Ala | Asp | Ala | 20 | 25 | 30 | |
| Ala | Leu | Gly | Glu | Val | Asn | Glu | Ser | Leu | Leu | Gln | Gln | Asp | Glu | Glu | Lys | 35 | 40 | 45 | |
| Ala | Leu | Tyr | Ala | Ala | Ala | Gln | Gly | Leu | Gln | Pro | Lys | Ile | Ala | Ala | Ala | 50 | 55 | 60 | |
| Val | Ala | Glu | Gly | Asn | Phe | Arg | Thr | Ala | Leu | Ser | Glu | Leu | Ala | Ser | Val | 65 | 70 | 75 | 80 |
| Lys | Pro | Gln | Val | Asp | Ala | Phe | Phe | Asp | Gly | Val | Met | Val | Met | Ala | Glu | 85 | 90 | 95 | |
| Asp | Ala | Ala | Val | Lys | Gln | Asn | Arg | Leu | Asn | Leu | Leu | Asn | Arg | Leu | Ala | 100 | 105 | 110 | |
| Glu | Gln | Met | Asn | Ala | Val | Ala | Asp | Ile | Ala | Leu | Leu | Gly | Glu | 115 | 120 | 125 | | | |

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 348 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

| | | | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|----|----|----|----|
| Met | Gln | Pro | Leu | Val | Ser | Val | Leu | Ile | Cys | Ala | Tyr | Asn | Val | Glu | Lys | 1 | 5 | 10 | 15 |
| Tyr | Phe | Ala | Gln | Ser | Leu | Ala | Ala | Val | Val | Asn | Gln | Thr | Trp | Arg | Asn | 20 | 25 | 30 | |

Leu Asp Ile Leu Ile Val Asp Asp Gly Ser Thr Asp Gly Thr Leu Ala
 35 40 45
 Ile Ala Lys Asp Phe Gln Lys Arg Asp Ser Arg Ile Lys Ile Leu Ala
 50 55 60
 Gln Ala Gln Asn Ser Gly Leu Ile Pro Ser Leu Asn Ile Gly Leu Asp
 65 70 75 80
 Glu Leu Ala Lys Ser Gly Gly Gly Gly Gly Glu Tyr Ile Ala Arg Thr
 85 90 95
 Asp Ala Asp Asp Ile Ala Ser Pro Gly Trp Ile Glu Lys Ile Val Gly
 100 105 110
 Glu Met Glu Lys Asp Arg Ser Ile Ile Ala Met Gly Ala Trp Leu Glu
 115 120 125
 Val Leu Ser Glu Glu Lys Asp Gly Asn Arg Leu Ala Arg His His Lys
 130 135 140
 His Gly Lys Ile Trp Lys Lys Pro Thr Arg His Glu Asp Ile Ala Ala
 145 150 155 160
 Phe Phe Pro Phe Gly Asn Pro Ile His Asn Asn Thr Met Ile Met Arg
 165 170 175
 Arg Ser Val Ile Asp Gly Gly Leu Arg Tyr Asp Thr Glu Arg Asp Trp
 180 185 190
 Ala Glu Asp Tyr Gln Phe Trp Tyr Asp Val Ser Lys Leu Gly Arg Leu
 195 200 205
 Ala Tyr Tyr Pro Glu Ala Leu Val Lys Tyr Arg Leu His Ala Asn Gln
 210 215 220
 Val Ser Ser Lys His Ser Val Arg Gln His Glu Ile Ala Gln Gly Ile
 225 230 235 240
 Gln Lys Thr Ala Arg Asn Asp Phe Leu Gln Ser Met Gly Phe Lys Thr
 245 250 255
 Arg Phe Asp Ser Leu Glu Tyr Arg Gln Thr Lys Ala Ala Ala Tyr Glu
 260 265 270
 Leu Pro Glu Lys Asp Leu Pro Glu Glu Asp Phe Glu Arg Ala Arg Arg
 275 280 285
 Phe Leu Tyr Gln Cys Phe Lys Arg Thr Asp Thr Pro Pro Ser Gly Ala
 290 295 300
 Trp Leu Asp Phe Ala Ala Asp Gly Arg Met Arg Arg Leu Phe Thr Leu
 305 310 315 320
 Arg Gln Tyr Phe Gly Ile Leu Tyr Arg Leu Ile Lys Asn Arg Arg Gln
 325 330 335
 Ala Arg Ser Asp Ser Ala Gly Lys Glu Gln Glu Ile
 340 345

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 306 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```

Met Asp Ile Val Phe Ala Ala Asp Asp Asn Tyr Ala Ala Tyr Leu Cys
 1           5           10           15
Val Ala Ala Lys Ser Val Glu Ala Ala His Pro Asp Thr Glu Ile Arg
 20           25           30
Phe His Val Leu Asp Ala Gly Ile Ser Glu Glu Asn Arg Ala Ala Val
 35           40           45
Ala Ala Asn Leu Arg Gly Gly Gly Asn Ile Arg Phe Ile Asp Val Asn
 50           55           60
Pro Glu Asp Phe Ala Gly Phe Pro Leu Asn Ile Arg His Ile Ser Ile
 65           70           75           80
Thr Thr Tyr Ala Arg Leu Lys Leu Gly Glu Tyr Ile Ala Asp Cys Asp
 85           90           95
Lys Val Leu Tyr Leu Asp Thr Asp Val Leu Val Arg Asp Gly Leu Lys
100           105           110
Pro Leu Trp Asp Thr Asp Leu Gly Gly Asn Trp Val Gly Ala Cys Ile
115           120           125
Asp Leu Phe Val Glu Arg Gln Glu Gly Tyr Lys Gln Lys Ile Gly Met
130           135           140
Ala Asp Gly Glu Tyr Tyr Phe Asn Ala Gly Val Leu Leu Ile Asn Leu
145           150           155           160
Lys Lys Trp Arg Arg His Asp Ile Phe Lys Met Ser Cys Glu Trp Val
165           170           175
Glu Gln Tyr Lys Asp Val Met Gln Tyr Gln Asp Gln Asp Ile Leu Asn
180           185           190
Gly Leu Phe Lys Gly Gly Val Cys Tyr Ala Asn Ser Arg Phe Asn Phe
195           200           205
Met Pro Thr Asn Tyr Ala Phe Met Ala Asn Gly Phe Ala Ser Arg His
210           215           220
Thr Asp Pro Leu Tyr Leu Asp Arg Thr Asn Thr Ala Met Pro Val Ala
225           230           235           240
Val Ser His Tyr Cys Gly Ser Ala Lys Pro Trp His Arg Asp Cys Thr
245           250           255
Val Trp Gly Ala Glu Arg Phe Thr Glu Leu Ala Gly Ser Leu Thr Thr
260           265           270

```

Val Pro Glu Glu Trp Arg Gly Lys Leu Ala Val Pro Pro Thr Lys Cys
 275 280 285
 Met Leu Gln Arg Trp Arg Lys Lys Leu Ser Ala Arg Phe Leu Arg Lys
 290 295 300
 Ile Tyr
 305

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 337 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Gln Pro Leu Val Ser Val Leu Ile Cys Ala Tyr Asn Ala Glu Lys
 1 5 10 15
 Tyr Phe Ala Gln Ser Leu Ala Ala Val Val Gly Gln Thr Trp Arg Asn
 20 25 30
 Leu Asp Ile Leu Ile Val Asp Asp Gly Ser Thr Asp Gly Thr Pro Ala
 35 40 45
 Ile Ala Arg His Phe Gln Glu Gln Asp Gly Arg Ile Arg Ile Ile Ser
 50 55 60
 Asn Pro Arg Asn Leu Gly Phe Ile Ala Ser Leu Asn Ile Gly Leu Asp
 65 70 75 80
 Glu Leu Ala Lys Ser Gly Gly Gly Glu Tyr Ile Ala Arg Thr Asp Ala
 85 90 95
 Asp Asp Ile Ala Ser Pro Gly Trp Ile Glu Lys Ile Val Gly Glu Met
 100 105 110
 Glu Lys Asp Arg Ser Ile Ile Ala Met Gly Ala Trp Leu Glu Val Leu
 115 120 125
 Ser Glu Glu Asn Asn Lys Ser Val Leu Ala Ala Ile Ala Arg Asn Gly
 130 135 140
 Ala Ile Trp Asp Lys Pro Thr Arg His Glu Asp Ile Val Ala Val Phe
 145 150 155 160
 Pro Phe Gly Asn Pro Ile His Asn Asn Thr Met Ile Met Arg Arg Ser
 165 170 175
 Val Ile Asp Gly Gly Leu Arg Phe Asp Pro Ala Tyr Ile His Ala Glu
 180 185 190
 Asp Tyr Lys Phe Trp Tyr Glu Ala Gly Lys Leu Gly Arg Leu Ala Tyr
 195 200 205
 Tyr Pro Glu Ala Leu Val Lys Tyr Arg Phe His Gln Asp Gln Thr Ser

| | | |
|---|-----|---------|
| 210 | 215 | 220 |
| Ser Lys Tyr Asn Leu Gln Gln Arg Arg Thr Ala Trp Lys Ile Lys Glu | | |
| 225 | 230 | 235 240 |
| Glu Ile Arg Ala Gly Tyr Trp Lys Ala Ala Gly Ile Ala Val Gly Ala | | |
| | 245 | 250 255 |
| Asp Cys Leu Asn Tyr Gly Leu Leu Lys Ser Thr Ala Tyr Ala Leu Tyr | | |
| | 260 | 265 270 |
| Glu Lys Ala Leu Ser Gly Gln Asp Ile Gly Cys Leu Arg Leu Phe Leu | | |
| | 275 | 280 285 |
| Tyr Glu Tyr Phe Leu Ser Leu Glu Lys Tyr Ser Leu Thr Asp Leu Leu | | |
| | 290 | 295 300 |
| Asp Phe Leu Thr Asp Arg Val Met Arg Lys Leu Phe Ala Ala Pro Gln | | |
| 305 | 310 | 315 320 |
| Tyr Arg Lys Ile Leu Lys Lys Met Leu Arg Pro Trp Lys Tyr Arg Ser | | |
| | 325 | 330 335 |
| Tyr | | |

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 280 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

| | |
|---|--|
| Met Gln Asn His Val Ile Ser Leu Ala Ser Ala Ala Glu Arg Arg Ala | |
| 1 5 10 15 | |
| His Ile Ala Asp Thr Phe Gly Ser Arg Gly Ile Pro Phe Gln Phe Phe | |
| 20 25 30 | |
| Asp Ala Leu Met Pro Ser Glu Arg Leu Glu Gln Ala Met Ala Glu Leu | |
| 35 40 45 | |
| Val Pro Gly Leu Ser Ala His Pro Tyr Leu Ser Gly Val Glu Lys Ala | |
| 50 55 60 | |
| Cys Phe Met Ser His Ala Val Leu Trp Glu Gln Ala Leu Asp Glu Gly | |
| 65 70 75 80 | |
| Leu Pro Tyr Ile Ala Val Phe Glu Asp Asp Val Leu Leu Gly Glu Gly | |
| 85 90 95 | |
| Ala Glu Gln Phe Leu Ala Glu Asp Thr Trp Leu Glu Glu Arg Phe Asp | |
| 100 105 110 | |
| Lys Asp Ser Ala Phe Ile Val Arg Leu Glu Thr Met Phe Ala Lys Val | |
| 115 120 125 | |

```

Ile Val Arg Pro Asp Lys Val Leu Asn Tyr Glu Asn Arg Ser Phe Pro
 130                      135                      140

Leu Leu Glu Ser Glu His Cys Gly Thr Ala Gly Tyr Ile Ile Ser Arg
145                      150                      155                      160

Glu Ala Met Arg Phe Phe Leu Asp Arg Phe Ala Val Leu Pro Pro Glu
                      165                      170                      175

Arg Ile Lys Ala Val Asp Leu Met Met Phe Thr Tyr Phe Phe Asp Lys
                      180                      185                      190

Glu Gly Met Pro Val Tyr Gln Val Ser Pro Ala Leu Cys Thr Gln Glu
195                      200                      205

Leu His Tyr Ala Lys Phe Leu Ser Gln Asn Ser Met Leu Gly Ser Asp
210                      215                      220

Leu Glu Lys Asp Arg Glu Gln Gly Arg Arg His Arg Arg Ser Leu Lys
225                      230                      235                      240

Val Met Phe Asp Leu Lys Arg Ala Leu Gly Lys Phe Gly Arg Glu Lys
                      245                      250                      255

Lys Lys Arg Met Glu Arg Gln Arg Gln Ala Glu Leu Glu Lys Val Tyr
260                      265                      270

Gly Arg Arg Val Ile Leu Phe Lys
275                      280

```

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 5859 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: unknown
- (ii) MOLECULE TYPE: DNA (genomic)
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: Neisseria gonorrhoeae
 - (B) STRAIN: F62
- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1491..2330
 - (C) GENE: lgtB
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```

CTGCAGGCCG TCGCCGTATT CAAACAACTG CCCGAAGCCG CCGCGCTCGC CGCCGCCAAC      60
AAACGCGTGC AAAACCTGCT GAAAAAAGCC GATGCCGCGT TGGGCGAAGT CAATGAAAGC      120
CTGCTGCAAC AGGACGAAGA AAAAGCCCTG TACGCTGCCG CGCAAGGTTT GCAGCCGAAA      180

```

| | | | | | | | |
|---|--|-------------|-------------|-------------|-------------------------|------------|-----|
| ATTGCCGCCG | CCGTCGCCGA | AGGCAATTTTC | CGAACC GCCT | TGTCCGAAC | T | GGCTTCCGTC | 240 |
| AAGCCGCAGG | TTGATGCCTT | CTTCGACGGC | GTGATGGTGA | TGGCGGAAGA | TGCCGCCGTA | 300 | |
| AAACAAAACC | GCCTGAACCT | GCTGAACCGC | TTGGCAGAGC | AGATGAACGC | GGTGGCCGAC | 360 | |
| ATCGCGCTTT | TGGGCGAGTA | ACCGTTGTAC | AGTCCAAATG | CCGTCTGAAG | CCTTCAGGCG | 420 | |
| GCATCAAATT | ATCGGGAGAG | TAAATTGCAG | CCTTTAGTCA | GCGTATTGAT | TTGCGCCTAC | 480 | |
| AACGTAGAAA | AATATTTTGC | CCAATCATTA | GCCGCCGTCG | TGAATCAGAC | TTGGCGCAAC | 540 | |
| TTGGATATTT | TGATTGTCGA | TGACGGCTCG | ACAGACGGCA | CACCTTGCCAT | TGCCAAGGAT | 600 | |
| TTTCAAAGC | GGGACAGCCG | TATCAAATC | CTTGACACAAG | CTCAAATTC | CGGCCTGATT | 660 | |
| CCCTCTTTAA | ACATCGGGCT | GGACGAATTG | GCAAAGTCGG | GGGGGGGGGG | GGGGGAATAT | 720 | |
| ATTGCGCGCA | CCGATGCCGA | CGATATTGCC | TCCCCCGGCT | GGATTGAGAA | AATCGTGGGC | 780 | |
| GAGATGGAAA | AAGACCGCAG | CATCATTGCG | ATGGGCGCGT | GGCTGGAAGT | TTTGTGCGAA | 840 | |
| GAAAAGGACG | GCAACCGGCT | GGCGCGGCAC | CACAAACACG | GCAAAATTTG | GAAAAAGCCG | 900 | |
| ACCCGGCACG | AAGACATCGC | CGCCTTTTTTC | CCTTTTCGGCA | ACCCCATACA | CAACAACACG | 960 | |
| ATGATTATGC | GGCGCAGCGT | CATTGACGGC | GGTTTTCGCTT | ACGACACCGA | GCGGGATTGG | 1020 | |
| GCGGAAGATT | ACCAATTTTG | GTACGATGTC | AGCAAATTGG | GCAGGCTGGC | TTATTATCCC | 1080 | |
| GAAGCCTTGG | TCAAATACCG | CCTTCACGCC | AATCAGGTTT | CATCCAAACA | CAGCGTCCGC | 1140 | |
| CAACACGAAA | TCGCGCAAGG | CATCCAAAAA | ACCGCCAGAA | ACGATTTTTT | GCAGTCTATG | 1200 | |
| GGTTTTAAAA | CCCGGTTTCGA | CAGCCTAGAA | TACCGCCAAA | CAAAAGCAGC | GGCGTATGAA | 1260 | |
| CTGCCGGAGA | AGGATTTGCC | GGAAGAAGAT | TTTGAACGCG | CCCGCCGGTT | TTTGTACCAA | 1320 | |
| TGCTTCAAAC | GGACGGACAC | GCCGCCCTCC | GGCGCGTGGC | TGGATTTTCGC | GGCAGACGGC | 1380 | |
| AGGATGAGGC | GGCTGTTTAC | CTTGAGGCAA | TACTTCGGCA | TTTGTACCG | GCTGATTAAA | 1440 | |
| AACCGCCGGC | AGGCGCGGTC | GGATTTCGGCA | GGGAAAGAAC | AGGAGATTTA | ATG CAA Met Gln 1 | 1496 | |
| AAC CAC GTT ATC AGC TTG GCT TCC GCC GCA GAA CGC AGG GCG CAC ATT | Asn His Val 5 | 5 | 10 | 15 | 1544 | | |
| GCC GCA ACC TTC GGC AGT CGC GGC ATC CCG TTC CAG TTT TTC GAC GCA | Ala Ala Thr 20 | 25 | 30 | 35 | 1592 | | |
| CTG ATG CCG TCT GAA AGG CTG GAA CGG GCA ATG GCG GAA CTC GTC CCC | Leu Met Pro Ser Glu Arg Leu Glu Arg Ala Met 35 | 40 | 45 | 50 | 1640 | | |
| GGC TTG TCG GCG CAC CCC TAT TTG AGC GGA GTG GAA AAA GCC TGC TTT | Gly Leu Ser Ala His 55 | 60 | 65 | 1688 | | | |

| | | | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|--|------|
| ATG | AGC | CAC | GCC | GTA | TTG | TGG | GAA | CAG | GCA | TTG | GAC | GAA | GGC | GTA | CCG | | 1736 |
| Met | Ser | His | Ala | Val | Leu | Trp | Glu | Gln | Ala | Leu | Asp | Glu | Gly | Val | Pro | | |
| | | | 70 | | | | | 75 | | | | | 80 | | | | |
| TAT | ATC | GCC | GTA | TTT | GAA | GAT | GAT | GTC | TTA | CTC | GGC | GAA | GGC | GCG | GAG | | 1784 |
| Tyr | Ile | Ala | Val | Phe | Glu | Asp | Asp | Val | Leu | Leu | Gly | Glu | Gly | Ala | Glu | | |
| | | 85 | | | | | 90 | | | | | 95 | | | | | |
| CAG | TTC | CTT | GCC | GAA | GAT | ACT | TGG | CTG | CAA | GAA | CGC | TTT | GAC | CCC | GAT | | 1832 |
| Gln | Phe | Leu | Ala | Glu | Asp | Thr | Trp | Leu | Gln | Glu | Arg | Phe | Asp | Pro | Asp | | |
| | 100 | | | | | 105 | | | | | 110 | | | | | | |
| TCC | GCC | TTT | GTC | GTC | CGC | TTG | GAA | ACG | ATG | TTT | ATG | CAC | GTC | CTG | ACC | | 1880 |
| Ser | Ala | Phe | Val | Val | Arg | Leu | Glu | Thr | Met | Phe | Met | His | Val | Leu | Thr | | |
| 115 | | | | | 120 | | | | | 125 | | | | | 130 | | |
| TCG | CCC | TCC | GGC | GTG | GCG | GAC | TAC | GGC | GGG | CGC | GCC | TTT | CCG | CTT | TTG | | 1928 |
| Ser | Pro | Ser | Gly | Val | Ala | Asp | Tyr | Gly | Gly | Arg | Ala | Phe | Pro | Leu | Leu | | |
| | | | 135 | | | | | 140 | | | | | | 145 | | | |
| GAA | AGC | GAA | CAC | TGC | GGG | ACG | GCG | GGC | TAT | ATT | ATT | TCC | CGA | AAG | GCG | | 1976 |
| Glu | Ser | Glu | His | Cys | Gly | Thr | Ala | Gly | Tyr | Ile | Ile | Ser | Arg | Lys | Ala | | |
| | | | 150 | | | | | 155 | | | | | 160 | | | | |
| ATG | CGT | TTT | TTC | TTG | GAC | AGG | TTT | GCC | GTT | TTG | CCG | CCC | GAA | CGC | CTG | | 2024 |
| Met | Arg | Phe | Phe | Leu | Asp | Arg | Phe | Ala | Val | Leu | Pro | Pro | Glu | Arg | Leu | | |
| | | 165 | | | | | 170 | | | | | 175 | | | | | |
| CAC | CCT | GTC | GAT | TTG | ATG | ATG | TTC | GGC | AAC | CCT | GAC | GAC | AGG | GAA | GGA | | 2072 |
| His | Pro | Val | Asp | Leu | Met | Met | Phe | Gly | Asn | Pro | Asp | Asp | Arg | Glu | Gly | | |
| | | 180 | | | | 185 | | | | | 190 | | | | | | |
| ATG | CCG | GTT | TGC | CAG | CTC | AAT | CCC | GCC | TTG | TGC | GCC | CAA | GAG | CTG | CAT | | 2120 |
| Met | Pro | Val | Cys | Gln | Leu | Asn | Pro | Ala | Leu | Cys | Ala | Gln | Glu | Leu | His | | |
| 195 | | | | | 200 | | | | | 205 | | | | | 210 | | |
| TAT | GCC | AAG | TTT | CAC | GAC | CAA | AAC | AGC | GCA | TTG | GGC | AGC | CTG | ATC | GAA | | 2168 |
| Tyr | Ala | Lys | Phe | His | Asp | Gln | Asn | Ser | Ala | Leu | Gly | Ser | Leu | Ile | Glu | | |
| | | | 215 | | | | | | 220 | | | | | 225 | | | |
| CAT | GAC | CGC | CGC | CTG | AAC | CGC | AAA | CAG | CAA | TGG | CGC | GAT | TCC | CCC | GCC | | 2216 |
| His | Asp | Arg | Arg | Leu | Asn | Arg | Lys | Gln | Gln | Trp | Arg | Asp | Ser | Pro | Ala | | |
| | | | 230 | | | | | 235 | | | | | 240 | | | | |
| AAC | ACA | TTC | AAA | CAC | CGC | CTG | ATC | CGC | GCC | TTG | ACC | AAA | ATC | GGC | AGG | | 2264 |
| Asn | Thr | Phe | Lys | His | Arg | Leu | Ile | Arg | Ala | Leu | Thr | Lys | Ile | Gly | Arg | | |
| | | 245 | | | | | 250 | | | | | 255 | | | | | |
| GAA | AGG | GAA | AAA | CGC | CGG | CAA | AGG | CGC | GAA | CAG | TTA | ATC | GGC | AAG | ATT | | 2312 |
| Glu | Arg | Glu | Lys | Arg | Arg | Gln | Arg | Arg | Glu | Gln | Leu | Ile | | | | | |

| | | | | | | |
|-------------|-------------|------------|-------------|------------|------------|------|
| CGGCTTCCCC | TTAAACATCA | GGCACATTTC | CATTACGACT | TATGCCCGCC | TGAAATTGGG | 2607 |
| CGAATACATT | GCCGATTGCG | ACAAAGTCCT | GTATCTGGAT | ACGGACGTAT | TGGTCAGGGA | 2667 |
| CGGCCTGAAG | CCCTTATGGG | ATACCGATT | GGGCGGTAAC | TGGGTCGGCG | CGTGCATCGA | 2727 |
| TTTGTGTTGTC | GAAAGGCAGG | AAGGATACAA | ACAAAAATC | GGTATGGCGG | ACGGAGAATA | 2787 |
| TTATTTCAAT | GCCGGCGTAT | TGCTGATCAA | CCTGAAAAAG | TGGCGGCGGC | ACGATATTTT | 2847 |
| CAAAATGTCC | TGCGAATGGG | TGGAACAATA | CAAGGACGTG | ATGCAATATC | AGGATCAGGA | 2907 |
| CATTTTGAAC | GGGCTGTTTA | AAGGCGGGGT | GTGTTATGCG | AACAGCCGTT | TCAACTTTAT | 2967 |
| GCCGACCAAT | TATGCCTTTA | TGGCGAACGG | GTTTGCGTCC | CGCCATACCG | ACCCGCTTTA | 3027 |
| CCTCGACCGT | ACCAATACGG | CGATGCCCCG | CGCCGTCAGC | CATTATTGCG | GCTCGGCAAA | 3087 |
| GCCGTGGCAC | AGGGACTGCA | CCGTTTGGGG | TGCGGAACGT | TTCACAGAGT | TGGCCGGCAG | 3147 |
| CCTGACGACC | GTTCCCGAAG | AATGGCGCGG | CAAACCTGCC | GTCCCGCCGA | CAAAGTGTAT | 3207 |
| GCTTCAAAGA | TGGCGCAAAA | AGCTGTCTGC | CAGATTCTTA | CGCAAGATTT | ATTGACGGGG | 3267 |
| CAGGCCGTCT | GAAGCCTTCA | GACGGCATCG | GACGTATCGG | AAAGGAGAAA | CGGATTGCAG | 3327 |
| CCTTTAGTCA | GCGTATTGAT | TTGCGCCTAC | AACGCAGAAA | AATATTTTGC | CCAATCATTG | 3387 |
| GCCGCCGTAG | TGGGGCAGAC | TTGGCGCAAC | TTGGATATTT | TGATTGTCGA | TGACGGCTCG | 3447 |
| ACGGACGGCA | CGCCCGCCAT | TGCCCGGCAT | TTCCAAGAAC | AGGACGGCAG | GATCAGGATA | 3507 |
| ATTTCCAATC | CCC GCAATTT | GGGCTTTATC | GCCTCTTTAA | ACATCGGGCT | GGACGAATTG | 3567 |
| GCAAAGTCGG | GGGGGGGGGA | ATATATTGCG | CGCACC GATG | CCGACGATAT | TGCCTCCCCC | 3627 |
| GGCTGGATTG | AGAAAATCGT | GGGCGAGATG | GAAAAAGACC | GCAGCATCAT | TGCGATGGGC | 3687 |
| GCGTGGTTGG | AAGTTTTGTC | GGAAGAAAAC | AATAAAAGCG | TGCTTGCCGC | CATTGCCCGA | 3747 |
| AACGGCGCAA | TTTGGGACAA | ACCGACCCGG | CATGAAGACA | TTGTCGCCGT | TTCCCTTTTC | 3807 |
| GGCAACCCCA | TACACAACAA | CACGATGATT | ATGAGGCGCA | GCGTCATTGA | CGGCGGTTTG | 3867 |
| CGGTTCGATC | CAGCCTATAT | CCACGCCGAA | GACTATAAGT | TTTGGTACGA | AGCCGGCAAA | 3927 |
| CTGGGCAGGC | TGGCTTATTA | TCCCGAAGCC | TTGGTCAAAT | ACCGCTTCCA | TCAAGACCAG | 3987 |
| ACTTCTTCCA | AATACAACCT | GCAACAGCGC | AGGACGGCGT | GGAAAATCAA | AGAAGAAATC | 4047 |
| AGGGCGGGGT | ATTGGAAGGC | GGCAGGCATA | GCCGTCGGGG | CGGACTGCCT | GAATTACGGG | 4107 |
| CTTTTGAAAT | CAACGGCATA | TGCGTTGTAC | GAAAAAGCCT | TGTCCGGACA | GGATATCGGA | 4167 |
| TGCCTCCGCC | TGTTCTTGTA | CGAATATTTT | TTGTCGTTGG | AAAAGTATTC | TTTGACCGAT | 4227 |
| TTGCTGGATT | TCTTGACAGA | CCGCGTGATG | AGGAAGCTGT | TTGCCGCACC | GCAATATAGG | 4287 |
| AAAATCCTGA | AAAAAATGTT | ACGCCCTTGG | AAATACCGCA | GCTATTGAAA | CCGAACAGGA | 4347 |
| TAAATCATGC | AAAACCACGT | TATCAGCTTG | GCTTCCGCCG | CAGAGCGCAG | GGCGCACATT | 4407 |

| | |
|--|------|
| GCCGATACCT TCGGCAGTCG CGGCATCCCG TTCCAGTTTT TCGACGCACT GATGCCGTCT | 4467 |
| GAAAGGCTGG AACAGGCGAT GGCGGAAGTC GTCCCCGGCT TGTCGGCGCA CCCCTATTTG | 4527 |
| AGCGGAGTGG AAAAAGCCTG CTTTATGAGC CACGCCGTAT TGTGGGAACA GGCCTTGGAT | 4587 |
| GAAGGTCTGC CGTATATCGC CGTATTTGAG GACGACGTTT TACTCGGCGA AGGCGCGGAG | 4647 |
| CAGTTCCTTG CCGAAGATAC TTGGTTGGAA GAGCGTTTTG ACAAGGATTC CGCCTTTATC | 4707 |
| GTCCGTTTGG AAACGATGTT TGCGAAAGTT ATTGTCAGAC CGGATAAAGT CCTGAATTAT | 4767 |
| GAAAACCGGT CATTTCCCTT GCTGGAGAGC GAACATTGTG GGACGGCTGG CTATATCATT | 4827 |
| TCGCGTGAGG CGATGCGGTT TTTCTTGGAC AGGTTTGCCG TTTTGCCGCC AGAGCGGATT | 4887 |
| AAAGCGGTAG ATTTGATGAT GTTTACTTAT TTCTTTGATA AGGAGGGGAT GCCTGTTTAT | 4947 |
| CAGGTTAGTC CCGCCTTATG TACCCAAGAA TTGCATTATG CCAAGTTTCT CAGTCAAAAC | 5007 |
| AGTATGTTGG GTAGCGATTT GGAAAAAGAT AGGGAACAAG GAAGAAGACA CCGCCGTTTCG | 5067 |
| TTGAAGGTGA TGTTTGACTT GAAGCGTGCT TTGGGTAAAT TCGGTAGGGA AAAGAAGAAA | 5127 |
| AGAATGGAGC GTCAAAGGCA GCGCGAGCTT GAGAAAGTTT ACGGCAGGCG GGTCATATTG | 5187 |
| TTCAAATAGT TTGTGTAAAA TATAGGGGAT TAAATCAGA AATGGACACA CTGTCATTCC | 5247 |
| CGCGCAGGCG GGAATCTAGG TCTTTAAACT TCGGTTTTTT CCGATAAATT CTTGCCGCAT | 5307 |
| TAAATTTCCA GATTCCCCTT TTCGCGGGGA TGACGGCGGG GGGATTGTTG CTTTTTCGGA | 5367 |
| TAAATCCCCT TGTTTTTTTCA TCTGCTAGGT AAAATCGCCC CAAAGCGTCT GCATCGCGGC | 5427 |
| GATGGCGGCG AGTGGGGCGG TTTCTGTGCG TAAATCCGT TTTCCGAGTG TAACCGCCTG | 5487 |
| AAAGCCGGCT TCAAATGCCT GTTGTCTTTC CTGTTCTGTC CAGCCGCCTT CGGGCCCCGAC | 5547 |
| CATAAAGACG ATTGCGCCCG ACGGTTGGCG GATGTCGCGG AGTTTGCAGG CGCGGTTGAT | 5607 |
| GCTCATAATC AGCTTGGTGT TTTCAGACGG CATTTTGTGCG AGTGCTTCAC GGTAGCCGAT | 5667 |
| GATGGGCAGT ACGGGGGGAA CGGTGTTTCT GCCGTTTGTG TCGCACGCGG AGATGACGAT | 5727 |
| TTCTTGCCAG CGTGCGAGGC GTTTGGCGGC GCGTTCTCCG TCGAGGCGGA CGATGCAGCG | 5787 |
| TTGCTGATG ACGGGCTGTA TGGCGGTTAC GCCGAGTTTC ACGCTTTTTT GCAGGGTGAA | 5847 |
| ATCCATGCGA TC | 5859 |

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 279 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Gln Asn His Val Ile Ser Leu Ala Ser Ala Ala Glu Arg Arg Ala
 1 5 10 15
 His Ile Ala Ala Thr Phe Gly Ser Arg Gly Ile Pro Phe Gln Phe Phe
 20 25 30
 Asp Ala Leu Met Pro Ser Glu Arg Leu Glu Arg Ala Met Ala Glu Leu
 35 40 45
 Val Pro Gly Leu Ser Ala His Pro Tyr Leu Ser Gly Val Glu Lys Ala
 50 55 60
 Cys Phe Met Ser His Ala Val Leu Trp Glu Gln Ala Leu Asp Glu Gly
 65 70 75 80
 Val Pro Tyr Ile Ala Val Phe Glu Asp Asp Val Leu Leu Gly Glu Gly
 85 90 95
 Ala Glu Gln Phe Leu Ala Glu Asp Thr Trp Leu Gln Glu Arg Phe Asp
 100 105 110
 Pro Asp Ser Ala Phe Val Val Arg Leu Glu Thr Met Phe Met His Val
 115 120 125
 Leu Thr Ser Pro Ser Gly Val Ala Asp Tyr Gly Gly Arg Ala Phe Pro
 130 135 140
 Leu Leu Glu Ser Glu His Cys Gly Thr Ala Gly Tyr Ile Ile Ser Arg
 145 150 155 160
 Lys Ala Met Arg Phe Phe Leu Asp Arg Phe Ala Val Leu Pro Pro Glu
 165 170 175
 Arg Leu His Pro Val Asp Leu Met Met Phe Gly Asn Pro Asp Asp Arg
 180 185 190
 Glu Gly Met Pro Val Cys Gln Leu Asn Pro Ala Leu Cys Ala Gln Glu
 195 200 205
 Leu His Tyr Ala Lys Phe His Asp Gln Asn Ser Ala Leu Gly Ser Leu
 210 215 220
 Ile Glu His Asp Arg Arg Leu Asn Arg Lys Gln Gln Trp Arg Asp Ser
 225 230 235 240
 Pro Ala Asn Thr Phe Lys His Arg Leu Ile Arg Ala Leu Thr Lys Ile
 245 250 255
 Gly Arg Glu Arg Glu Lys Arg Arg Gln Arg Arg Glu Gln Leu Ile Gly
 260 265 270
 Lys Ile Ile Val Pro Phe Gln
 275

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 21 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: PCR primer
- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

GCCGAGAAAA CTATTGGTGG A

21

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 22 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: unknown

- (ii) MOLECULE TYPE: cDNA
- (iii) HYPOTHETICAL: NO
- (iv) ANTI-SENSE: NO
- (vi) ORIGINAL SOURCE:
 - (A) ORGANISM: PCR primer

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

AAAACATGCA GGAATTGACG AT

22

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 348 amino acids
 - (B) TYPE: amino acid
 - (D) TOPOLOGY: linear

- (ii) MOLECULE TYPE: protein

- (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Leu | Gln | Pro | Leu | Val | Ser | Val | Leu | Ile | Cys | Ala | Tyr | Asn | Val | Glu | Lys |
| 1 | | | | 5 | | | | | 10 | | | | | 15 | |
| Tyr | Phe | Ala | Gln | Ser | Leu | Ala | Ala | Val | Val | Asn | Gln | Thr | Trp | Arg | Asn |
| | | 20 | | | | | | 25 | | | | | 30 | | |
| Leu | Asp | Ile | Leu | Ile | Val | Asp | Asp | Gly | Ser | Thr | Asp | Gly | Thr | Leu | Ala |
| | | 35 | | | | | 40 | | | | | 45 | | | |
| Ile | Ala | Lys | Asp | Phe | Gln | Lys | Arg | Asp | Ser | Arg | Ile | Lys | Ile | Leu | Ala |
| | | 50 | | | | 55 | | | | | 60 | | | | |
| Gln | Ala | Gln | Asn | Ser | Gly | Leu | Ile | Pro | Ser | Leu | Asn | Ile | Gly | Leu | Asp |
| 65 | | | | | 70 | | | | 75 | | | | | 80 | |

(2) INFORMATION FOR SEQ ID NO:12:

(A) LENGTH: 337 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Leu Gln Pro Leu Val Ser Val Leu Ile Cys Ala Tyr Asn Ala Glu Lys
 1 5 10 15
 Tyr Phe Ala Gln Ser Leu Ala Ala Val Val Gly Gln Thr Trp Arg Asn
 20 25 30
 Leu Asp Ile Leu Ile Val Asp Asp Gly Ser Thr Asp Gly Thr Pro Ala
 35 40 45
 Ile Ala Arg His Phe Gln Glu Gln Asp Gly Arg Ile Arg Ile Ile Ser
 50 55 60
 Asn Pro Arg Asn Leu Gly Phe Ile Ala Ser Leu Asn Ile Gly Leu Asp
 65 70 75 80
 Glu Leu Ala Lys Ser Gly Gly Gly Glu Tyr Ile Ala Arg Thr Asp Ala
 85 90 95
 Asp Asp Ile Ala Ser Pro Gly Trp Ile Glu Lys Ile Val Gly Glu Met
 100 105 110
 Glu Lys Asp Arg Ser Ile Ile Ala Met Gly Ala Trp Leu Glu Val Leu
 115 120 125
 Ser Glu Glu Asn Asn Lys Ser Val Leu Ala Ala Ile Ala Arg Asn Gly
 130 135 140
 Ala Ile Trp Asp Lys Pro Thr Arg His Glu Asp Ile Val Ala Val Phe
 145 150 155 160
 Pro Phe Gly Asn Pro Ile His Asn Asn Thr Met Ile Met Arg Arg Ser
 165 170 175
 Val Ile Asp Gly Gly Leu Arg Phe Asp Pro Ala Tyr Ile His Ala Glu
 180 185 190
 Asp Tyr Lys Phe Trp Tyr Glu Ala Gly Lys Leu Gly Arg Leu Ala Tyr
 195 200 205
 Tyr Pro Glu Ala Leu Val Lys Tyr Arg Phe His Gln Asp Gln Thr Ser
 210 215 220
 Ser Lys Tyr Asn Leu Gln Gln Arg Arg Thr Ala Trp Lys Ile Lys Glu
 225 230 235 240
 Glu Ile Arg Ala Gly Tyr Trp Lys Ala Ala Gly Ile Ala Val Gly Ala
 245 250 255
 Asp Cys Leu Asn Tyr Gly Leu Leu Lys Ser Thr Ala Tyr Ala Leu Tyr
 260 265 270
 Glu Lys Ala Leu Ser Gly Gln Asp Ile Gly Cys Leu Arg Leu Phe Leu
 275 280 285
 Tyr Glu Tyr Phe Leu Ser Leu Glu Lys Tyr Ser Leu Thr Asp Leu Leu
 290 295 300
 Asp Phe Leu Thr Asp Arg Val Met Arg Lys Leu Phe Ala Ala Pro Gln
 305 310 315 320
 Tyr Arg Lys Ile Leu Lys Lys Met Leu Arg Pro Trp Lys Tyr Arg Ser
 325 330 335

